REMARKS

Claims 1-21 are pending in the application.

Claim 3 is original and claims 1, 2, and 4-21 have been previously presented. Applicants respectfully submit that the present claims are in condition for allowance or appeal. Applicants respectfully request that the Examiner reconsider the present claims in view of the following discussion.

Claims 1-21 stand rejected under 35 U.S.C. §103(a) as unpatentable over Bahl et al., "A Tree-Based Statistical Language Model for Natural Language Speech Recognition", hereinafter, Bahl, in view of U.S. Patent No. 6,292,772 to Kantrowitz.

Applicants respectfully traverse the rejection based on the following discussion. The following paragraphs are numbered for ease of future reference.

I. The 35 U.S.C. 103(a) Rejection over Bahl and Kantrowitz

A. The Bahl Disclosure

[0001] It is a fact that Bahl discloses, "This paper is concerned with the problem of 'predicting' the next word a speaker will say, given the words already spoken; specifically, the problem is to estimate the probability that a given word will be the next word uttered." (Abstract, first sentence).

[0002] It is a fact that Bahl discloses, "In this model [i.e., the *N*-gram language model], word sequences are treated as equivalent if and only if they end with the same *N*-1 words. Typically, *N*=3, in which case the model is referred to as a 3-gram or trigam model. The trigam model is based upon the approximation $\Pr\{w_i|w_1, w_2, \ldots w_{i-1}\} \approx \Pr\{w_i|w_{i-2},w_{i-1}\}$, which is clearly inexact, but apparently quite useful." (page 1001, col. 2, which is cited by the Final Action).

[0003] It is a fact that Bahl discloses, "Maximum-likelihood estimates of *N*-gram probabilities can be obtained from their relative frequencies in a large body of training text. But since many legitimate *N*-grams are likely to be missing from the training text, it is necessary to 'smooth' the maximum-likelihood estimates so as to avoid probabilities of zero." (page 1001, col. 2, which is cited by the Final Action).

[0004] It is a fact that Bahl discloses, "We tested the foregoing ideas on tree-based language models in a pilot experiment involving a 5000-word vocabulary. The vocabulary consisted of the 5000 most frequent words in a database of IBM office correspondence." (page 1005, III. Results, first paragraph).

[0005] It is a fact that Bahl discloses, "The training and test data were drawn from 550 books and magazines which ranged from intellectually stimulating romantic novels to silly issues of *Datamation*." (page 1005, III. Results, second paragraph, first sentence).

[0006] It is a fact that Bahl discloses, "No book or magazine was split between one of the above categories. Words not in the vocabulary were treated as a single generic "unknown" word." (Page 1006, first paragraph, which is cited by the Final Action).

[0007] It is a fact that Bahl discloses, "The purpose of the experiment was to predict the 21st word of a 21-gram, given the first 20 words." (Page 1006, second paragraph, first sentence, where the second paragraph is cited by the Final Action).

[0008] It is a fact that Bahl discloses, "We constructed a tree with pylonic questions using the tree-growing and set-construction algorithms of Section II, with two minor changes." (Page 1006, third paragraph, first sentence, where the third paragraph is cited by the Final Action).

B. The Kantrowitz Disclosure

[0009] It is a fact that Kantrowitz discloses, "The present invention is different from these systems in that it identifies the language of individual words with very high accuracy, not entire documents. This allows the present invention to operate on a word-by-word basis, correctly identifying the language of words even when the document contains multiple languages (e.g., Canadian parliamentary proceedings contain both English and French) or includes short quotes of one language within a document that is mostly another language." (col. 2, lines 27-35).

[0010] It is a fact that Kantrowitz discloses, "This allows language-specific functionality, such as language-specific spelling correction and transliteration (e.g., ASCII-to-Kanji conversion of Japanese Romanji to Kanji letters) to occur on a word-by-word basis. The language identification statistics for the individual words of a document can be combined to identify the

overall language of the document with a much higher cumulative accuracy than the state of the art. It can also identify the number of languages present in mixed-language documents, the identity of the language and the relative frequency of the occurrence of the language's lexicon." (col. 2, lines 35-45).

C. Arguments

[0011] It is a fact that the Specification discloses, "The existing techniques described above [i.e., *N*-gram model; decision, classification, and regression trees; context-free grammars] are not entirely adequate in processing mixed language expressions, which arise, for example, in spoken language. As an example, English language words and phrases are often embedded in a speaker's native language, due to the dominance of English as an international language." (Page 2, line 31 to page 3, line1).

[0012] It is a fact that the Specification discloses, "For example, many Indians use English words embedded in Hindi sentences during conversation." (Page 3, lines 23-24).

[0013] It is a fact that the Specification discloses, "The monolingual language generator uses a mixed-language word history to generate a monolingual word history. The monolingual history is in turn used by a monolingual language model. A resulting next-word hypothesis is used by a next-word language change model, which uses word equivalence probabilities to convert the next word in the monolingual word hypothesis to the next word in the foreign language." (Page 3, lines 11-15).

[0014] It is a fact that the Specification discloses, "This mixed language sentence is 'Delhi becomes very GARM in summer'. In this sentence, 'GARM' is a Hindi word embedded in an otherwise English language sentence. Now, during speech recognition of this sentence, to compute the language model probability of the word 'GARM', a mixed language model between Hindi and English would ordinarily be required. As described, such a model is not available, as the text data for this kind of usage is not available." (Page 12, lines 28-33).

[0015] It is a fact that the Specification discloses, "Instead, the word equivalence probabilities of 'GARM' with the equivalent English words (such as 'hot', 'warm', 'boiled', 'temperature', etc.) These equivalent probabilities are estimated by a parallel text corpus between

Hindi and English as described." (Page 13, lines 1-4).

[0016] With respect to independent claims 1, 8, and 9, the Final Action states, "Bahl teaches storing word equivalence probabilities resulting from words of a first language and words in at least one other language (Page 1001, Col. 2)." (Final Action, page 13, lines 1-3).

[0017] Applicants respectfully submit that the Final Action has misconstrued Bahl, because Bahl is concerned with the problem of estimating the probability that a given word will be the next word uttered in a sequence of words from but a single language (Abstract, first sentence).

[0018] As is shown by the selection of training and test data, drawn from 550 books and magazines which ranged from intellectually stimulating romantic novels to silly issues of *Datamation* (Page 2, line 31 to page 3, line1), Bahl's training data and test data were both in the English language. (emphasis added).

[0019] In contrast, the present invention clearly describes at least the feature of "storing word equivalence probabilities resulting from words of a first language and words in at least one other language", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9. Nowhere does Bahl disclose, teach or suggest storing word equivalence probabilities, i.e., word equivalence probabilities, which convert the next word in the monolingual word hypothesis (i.e., in a first language) to the next word in the foreign language (in at least one other language) (Specification, page 3, lines 11-15). Instead, Bahl merely discloses the problem of estimating the probability that a given word will be the next word uttered in a sequence of words from but a single language (Abstract, first sentence).

[0020] With respect to independent claims 1, 8, and 9, the Final Action states that Bahl teaches "generating a monolingual word history and using the stored word equivalence probabilities, wherein said mixed language word history comprises words in said first language and words in said at least one other language, and wherein said mixed language word history and said monolingual word history each comprise a history of previous words in a sentence-based word sequence (Page 1001, Col. 2)". (Final Action, page 13, second paragraph).

[0021] As argued above, Bahl does not disclose, teach or suggest the present invention's features of: storing word equivalence probabilities, i.e., word equivalence probabilities, which

convert the next word in the monolingual word hypothesis (i.e., in a first language) to the next word in the foreign language (in at least one other language) (Specification, page 3, lines 11-15). Instead, Bahl merely discloses the problem of estimating the probability that a given word will be the next word uttered in a sequence of words from but a single language (Abstract, first sentence).

[0022] Furthermore, Bahl does not disclose, teach or suggest at least the feature of a: "mixed language word history compris[ing] words in said first language and words in said at least one other language", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9, because Bahl merely discloses the problem of estimating the probability that a given word will be the next word uttered in a sequence of words from but a single language (Abstract, first sentence). Nowhere does Bahl disclose, teach or suggest the concept of a mixed language word history comprising words in a first language and words in at least one other language, as clearly described in independent claims 1, 8, and 9.

[0023] In addition, Bahl does not disclose, teach or suggest at least the feature of a "mixed language word history and said monolingual word history each compris[ing] a history of previous words in a sentence-based word sequence", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9, because Bahl merely discloses the problem of estimating the probability that a given word will be the next word uttered in a sequence of words from but a single language (Abstract, first sentence). Nowhere does Bahl disclose, teach or suggest the concept of a mixed language word history comprising a history of previous words in a sentence-based word sequence, as clearly described in independent claims 1, 8, and 9.

[0024] With respect to independent claims 1, 8, and 9, the Final Action states that Bahl teaches "determining a probability of a next word (Page 1002 Col. 2) in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities (Page 1001 Col. 2), wherein said probability of said next word predicts a next word in said mixed language expression (Page 1006 Col. 1 paragraphs 1-3).

[0025] As similarly argued above, Bahl does not disclose, teach or suggest "determining a probability of a next word (Page 1002 Col. 2) in a mixed language expression", as recited in claims 1 and 8, and as similarly recited in claim 9, because Bahl merely discloses the problem of

estimating the probability that a given word will be the next word uttered in a sequence of words from but a single language (Abstract, first sentence).

[0026] As argued above, Bahl does not disclose, teach or suggest storing word equivalence probabilities, i.e., word equivalence probabilities, which convert the next word in the monolingual word hypothesis (i.e., in a first language) to the next word in the foreign language (in at least one other language) (Specification, page 3, lines 11-15). Hence, Bahl cannot "determin[e] a probability of a next word in a mixed language expression based upon … the stored word equivalence probabilities", as clearly described in independent claims 1, 8, and 9.

[0027] Furthermore, Bahl cannot disclose, teach or suggest at least the present invention's feature of: "wherein said probability of said next word predicts a next word in said mixed language expression", because Bahl merely discloses the problem of estimating the probability that a given word will be the next word uttered in a sequence of words from but a single language (Abstract, first sentence).

[0028] In addition, the recitation of page 1006, the first three paragraphs of Bahl by the Final Action clearly shows the Final Action's misconstruction of Bahl, because in these three paragraphs, Bahl is discussing a tree-growing algorithm in which the set S^c_i minimizes the average conditional entropy at the current node (page 1003, col. 2, second paragraph), where determining S^c_i amounts to partitioning the values taken by X_i [a discrete random variable] into two groups: those in S^c_i and those not in S^c_i (page 1004, col. 1, second paragraph), and finally, constructing a tree using these tree-growing algorithms (page 1006, col. 1, third paragraph). Nowhere does Bahl disclose, teach or suggest the present invention's features of: "determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities, wherein said probability of said next word predicts a next word in said mixed language expression", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9, and as alleged by the Final Action. Instead, Bahl is discussing partitioning nodes of a tree based on entropy.

[0029] For at least the reasons outlined above, Applicants respectfully submit that Bahl does not disclose, teach or suggest at least the present invention's features of: "storing word

equivalence probabilities relating to words of a first language and words in at least one other language; generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities, wherein said mixed language word history comprises words in said first language and words in said at least one other language, and wherein said mixed language word history and said monolingual word history each comprise a history of previous words in a sentence-based word sequence; … and determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities, wherein said probability of said next word predicts a next word in said mixed language expression", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9.

[0030] With respect to independent claims 1, 8, and 9, the Final Action states, "However, Bahl fails to teach a method for language modeling of mixed language expressions". (Final Action, page 13, next to last paragraph). The Final Action then cites Kantrowitz, col. 6, lines 7-64, for presumably teaching a method for language modeling of mixed language expressions. (Final Action, page 13, next to last paragraph).

[0031] Applicants respectfully submit that the Final Action has misconstrued the teachings of Kantrowitz.

[0032] Applicants further respectfully submit that Kantowitz does not disclose, teach or suggest at least the present invention's features of: determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities, wherein said probability of said next word predicts a next word in said mixed language expression", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9.

[0033] Instead, Kantrowitz discloses, "The invention herein goes beyond the state of the art by being able to identify the language of individual words in isolation with high accuracy. The accuracy in identifying the language of individual words typically is equal to that of whole-document language identification systems. When the language identification of individual words is combined for all the words in a document, the overall accuracy significantly exceeds that of

whole-document systems. Moreover, the ability to identify the language of individual words permits document processing resources to be applied on a word-by-word basis." (col. 6, lines 41-51, which is cited by the Final Action).

[0034] Applicants respectfully submit that Kantrowitz does not determine a probability of a next word in a mixed language expression, as clearly described by independent claims 1, 8, and 9 of the present invention. Instead, Kantrowitz's "invention consists of a computer method for identifying words of a particular language. As used herein, term "word" is used in its normal sense to mean a string of characters that as ordered have meaning in a given language." (col. 3, lines 9-12). Kantrowitz also discloses that "The matching expression tests a string of characters for an n-gram match at the beginning of the word, followed by one or more of a small set of n-grams within the word, followed by a match at the end of the word." (col. 3, lines 16-20). Thus, Kantowitz discloses a method of identifying a single word in a particular language by identifying a string of characters analyzed by an n-gram model (n-grams may apply to either characters or words, or even phrases). Kantrowtiz does not determine a probability of a next word in a mixed language expression, as described by independent claims 1, 8, and 9, because Kantrowitz is not modeling language expressions of multiple words using a mixed language; instead, Kantrowitz is identifying but a single word (in isolation from other words) in a particular language by identifying characters which constitute the single word.

[0035] Kantrowitz further discloses that "[t]he method of recognizing the language of a single word has applications to spelling and grammar correction (e.g., identifying the appropriate language resources on a document, paragraph, sentence or even individual word basis), the automatic invocation of transliteration software based on the language of the words (e.g., automatic ASCII to Kanji substitution without requiring the user to explicitly switch into a Kanji mode), etc." (col. 6, lines 26-33). Again, Kantrowitz is identifying but a single word in a particular language to which further applications may apply.

[0036] Nowhere does Kantrowitz disclose, teach or suggest at least the present invention's features of: determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities, wherein said probability of said next word predicts a next word in said mixed

language expression", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9, because Kantrowitz is merely identifying the language of individual words in isolation by analyzing the characters of the word (col. 6, lines 42-43).

[0037] Furthermore, Kantrowitz does not cure the deficiencies of Bahl argued above.

[0038] Nowhere does Kantrowitz disclose, teach or suggest at least the present invention's features of: "storing word equivalence probabilities relating to words of a first language and words in at least one other language; generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities, wherein said mixed language word history comprises words in said first language and words in said at least one other language, and wherein said mixed language word history and said monolingual word history each comprise a history of previous words in a sentence-based word sequence; … and determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities, wherein said probability of said next word predicts a next word in said mixed language expression", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9.

[0039] Instead, Kantrowitz merely identifies the language of individual words in isolation by analyzing the characters of the word (col. 6, lines 42-43).

[0040] For at least the reasons outlined above, Applicants respectfully submit that Bahl and Kantrowitz, either individually or in combination, do not disclose, teach or suggest at least the present invention's features of: "storing word equivalence probabilities relating to words of a first language and words in at least one other language; generating a monolingual word history in the first language based upon a mixed language word history and using the stored word equivalence probabilities, wherein said mixed language word history comprises words in said first language and words in said at least one other language, and wherein said mixed language word history and said monolingual word history each comprise a history of previous words in a sentence-based word sequence; … and determining a probability of a next word in a mixed language expression based upon the monolingual next word hypothesis probabilities and the stored word equivalence probabilities, wherein said probability of said next word predicts a next

word in said mixed language expression", as recited in independent claims 1 and 8, and as similarly recited in independent claim 9. Accordingly, Bahl and Kantrowitz, either individually or in combination, fail to render obvious the subject matter of previously presented, independent claims 1, 8, and 9, and dependent claims 2-7 and 10-21 under 35 U.S.C. §103(a). Withdrawal of the rejection of claims 1-21 under 35 U.S.C. §103(a) as unpatentable over Bahl and Kantrowitz is respectfully solicited.

II. Formal Matters and Conclusion

Claims 1-21 are pending in the application.

With respect to the rejection of the claims over the cited prior art, Applicants respectfully argue that the present claims are distinguishable over the prior art of record. In view of the foregoing, the Examiner is respectfully requested to reconsider and withdraw the rejections to the claims.

In view of the foregoing, Applicants submit that claims 1-21, all the claims presently pending in the application, are in condition for allowance. The Examiner is respectfully requested to pass the above application to issue at the earliest time possible.

Should the Examiner find the application to be other than in condition for allowance, the Examiner is requested to contact the undersigned at the local telephone number listed below to discuss any other changes deemed necessary.

Please charge any deficiencies and credit any overpayments to Attorney's Deposit Account Number 09-0441.

Respectfully submitted,

Dated: December 23, 2008

/Peter A. Balnave/
Peter A. Balnave, Ph.D.
Registration No. 46,199

Gibb I. P. Law Firm, LLC
2568-A Riva Road, Suite 304
Annapolis, MD 21401
Voice: (410) 573-5255
Fax: (301) 261-8825
Email: Balnave@GibbIPLaw.com
Customer Number: 29154